

Qy 421 Thr***AlaGluGlyGluAla*****LeuAspLeuArgGly***HisPheGlnLeuLeu 440
 ||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1261 ACAGGGGCTGAAGGGAAAGCAAGGCCTTGTGATCTTAGGGACAACTTTCAACTCTC 1320

 Qy 441 ProPheGlySerGlyArg***MetCysProGlyVal***LeuAlaThrSerGly***Ala 460
 ||||| ||||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1321 CCATTGGGTCTGGAGGAGAATGTGCCCTGGAGTCAATCTGGCTACTTCGGGAATGGCA 1380

 Qy 461 ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGln 480
 ||||| ||||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1381 ACACTTCTTGCATCTCTTATTCACTGCTTGACTTGCAAGTGCTGGTCCACAAGGACAG 1440

 Qy 481 IleLeuLysGly***AspAlaLysValSerMetGluGluArgAlaGlyLeuThrValPro 500
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1441 ATATTGAAGGGTGGTGACGCCAAAGTTAGCATGGAAGAGAGAGGCCGGCTCACTGTTCCA 1500

 Qy 501 ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu 520
 ||||| ||||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1501 AGGGCACATAGTCTTGTCTGTGTTCCACTTGCAAGGATCGGCCTGCATCTAAACTCCTT 1560

 Qy 521 Ser 521
 |||
 Db 1561 TCT 1563

RESULT 9

AF135484

LOCUS AF135484 1722 bp mRNA linear PLN 02-AUG-1999
DEFINITION Glycine max cytochrome P450 monooxygenase CYP93C1v2p (CYP93C1v2)
 mRNA, complete cds.
ACCESSION AF135484
VERSION AF135484.1 GI:5059123
KEYWORDS
SOURCE Glycine max.
ORGANISM Glycine max.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
REFERENCE 1 (bases 1 to 1722)
AUTHORS Steele,C.L., Gijzen,M., Qutob,D. and Dixon,R.A.
TITLE Molecular characterization of the enzyme catalyzing the aryl
 migration reaction of isoflavanoid biosynthesis in soybean
JOURNAL Arch. Biochem. Biophys. 367 (1), 146-150 (1999)
MEDLINE 99306846
PUBMED 10375412
REFERENCE 2 (bases 1 to 1722)
AUTHORS Steele,C.L., Gijzen,M., Qutob,D. and Dixon,R.A.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-1999) Plant Biology, Noble Foundation, 2510 Sam
 Noble Pkwy, Ardmore, OK 73402, USA
FEATURES Location/Qualifiers
source 1. .1722
 /organism="Glycine max"
 /db_xref="taxon:3847"
gene 1. .1722
 /gene="CYP93C1v2"
CDS 36. .1601

```

/gene="CYP93C1v2"
/function="2-hydroxyisoflavanone synthase"
/note="The functional expression of this cDNA demonstrated
2-hydroxyisoflavanone synthase activity"
/codon_start=1
/product="cytochrome P450 monooxygenase CYP93C1v2p"
/protein_id="AAD38929.1"
/db_xref="GI:5059124"
/translation="MLLELALGLLVLLALFLHLRPTPTAKSKALRHLPPSPKPRLPF
IGHLHLLDKKLHYALIDLSKKGPLFSLYFGSMPTVVASTPELFKLFLQTHEATSFN
TRFQTSAIRRLTYDSSVAMVPFGPYWKFRKLMNDLLNATTVNKLRLPRTQQIRKFL 160
RVMAQAEAQKPLDTEELLKWTNSTISMMMLGEAEEIRDIAREVLIKIFGEYSLTDFI
WPLKHLKVGKYEKRIDDILNKFDPPVERVIKKRRELVRRRNKEVVEGEVSGVFLDTL
LEFAEDETMIEIKITKDHIKGLVVDFPSAGTDSTAVATEWALAEINNNPKVLEKAREEV
YSVVGKDRLVDEYDTQNLPIRAIVKETFRMHPPLPVVKRKCTEECEINGYVIEGAL
ILFNYWQVGRDPKYWDRPSEFRPERFLETGAEGEAGPLDLRGQHFQLLPFGSGRRMCP
GVNLATSGMATLLASLIQCFDLQVLGPQGQILKGGDAKVSMEERAGLTVPRAHSLVCV
PLARIGVASKLLS"

misc_difference 445
/gene="CYP93C1v2"
/note="difference compared to GenBank Accession Number
AF022462; causes change of P to L (aa 140)"
/replace="c"
misc_difference 502
/gene="CYP93C1v2"
/note="difference compared to GenBank Accession Number
AF022462; causes change of T to I (aa 156)"
/replace="c"
misc_difference 918
/gene="CYP93C1v2"
/note="difference compared to GenBank Accession Number
AF022462; causes change of E to K (aa 295)"
/replace="g"
misc_difference 1612..1620
/gene="CYP93C1v2"
/note="9 bp insertion compared to GenBank Accession Number
AF022462"
/replace=""

```

BASE COUNT 470 a 422 c 407 g 423 t
ORIGIN

Alignment Scores:

Pred. No.:	1.52e-259	Length:	1722
Score:	2267.00	Matches:	454
Percent Similarity:	87.14%	Conservative:	0
Best Local Similarity:	87.14%	Mismatches:	67
Query Match:	94.62%	Indels:	0
DB:	8	Gaps:	0

US-09-857-581-66 (1-521) x AF135484 (1-1722)

Qy	1 MetLeuLeuGluLeuAlaLeuGlyLeu***ValLeuAlaLeuPhe***HisLeuArgPro 20	Leu	Leu
Db	36 ATGTTGCTTGAACTTGCACCTGGTTATTGGTTTGCGCTCTGTTCTGCACCTGCGTCCC 95		
Qy	21 ThrPro***Ala***SerLysAlaLeuArgHisLeuProAsnProProSerPro***Pro 40		

Db 96 ACACCCACTGCAAAATCAAAGCACTTCGCCATCTCCAAACCCACCAAGCCAAAGCCT 155
 Qy 41 ArgLeuProPheIleGlyHis***HisLeuLeuLysAspLysLeuLeuHisTyrAla*** 60
 Db 156 CGTCTTCCCTTCATAGGACACCTTCATCTCTAAAAGACAAACTTCTCCACTACGCAC 215
 Qy 61 IleAspLeuSerLysLysHisGlyProLeuPheSer*****pheGlySerMetProThr 80
 Db 216 ATCGACCTCTCCAAAAAACATGGTCCCTATTCTCTACTTTGGCTCCATGCCAACC 275
 Qy 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln*****GluAlaThrSer 100
 Db 276 GTTGTGCTCCACACCAGAATTGTTCAAGCTCTCCAAACGCACGGCAACTTCC 335
 Qy 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla 120
 Db 336 TTCAACACAAGGTTCAAACCTCAGCCATAAGACGCCACCTATGATAGCTCAGTGGCC 395
 Qy 121 *****Pro***GlyProTyrTrp***PheValArgLysLeuIleMetAsnAspLeuLeu 140
 Db 396 ATGGTTCCCTCGGACCTACTGGAAGTTCGTGAGGAAGCTCATCATGAACGACCTCTC 455
 Qy 141 AsnAlaThrThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgLys***Leu 160
 Db 456 AACGCCACCACTGTAAACAAGTTGAGGCCTTGAGGACCCAACAGATCCGCAAGTTCTC 515
 Qy 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluGluLeuLeu 180
 Db 516 AGGGTTATGGCCAAGGCGCAGAGGCACAGAAGCCCCTTGACTTGACCGAGGAGCTCTG 575
 Qy 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyGluAlaGluGluIleArgAsp 200
 Db 576 AAATGGACCAACAGCACCATCTCCATGATGATGCTCGCGAGGCTGAGGAGATCAGAGAC 635
 Qy 201 IleAlaArgGluValLeuLysIle***GlyGluTyrSerLeuThrAspPheIle***Pro 220
 Db 636 ATCGCTCGCGAGGTTCTTAAGATCTTGCGAATACAGCCTCACTGACTTCATCTGGCCA 695
 Qy 221 LeuLys***LeuLysValGlyLysTyrGluLysArgIleAspAspIleLeuAsnLysPhe 240
 Db 696 TTGAAGCATCTCAAGGTTGAAAGTATGAGAAGAGGATCGACGACATCTGAACAAGTTC 755
 Qy 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
 Db 756 GACCCTGTCGTTGAAAGGGTCATCAAGAACGCGCGTGGAGATCGAGACATCTGAACAAGAAC 815
 Qy 261 GlyGlu*****GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAla 280
 Db 816 GGAGAGGTTGTTGAGGGTGAGGTCAAGCGGGTTTCTTGACACTTGCTTGAATTGCT 875
 Qy 281 GluAspGluThr***GluIleLysIleThrLys***IleLysGlyLeuValValAsp 300
 Db 876 GAGGATGAGACCATGGAGATCAAATCACCAAGGACACATCAAGGGTCTGTTGTCGAC 935
 Qy 301 ***PheSerAlaGly***AspSerThrAla*****ThrGluTrpAlaLeuAlaGluLeu 320
 Db 936 TTTTCTCGGCAGGAACAGACTCCACAGCGGTGGCAACAGAGTGGCATTGGCAGAAC 995

Qy	321	IleAsnAsnPro***ValLeu*****AlaArgGluGlu***TyrSerValValGlyLys	340
Db	996	ATCAACAATCTAAGGTGGAAAAGGCTCGTGAGGGTCTACAGTGTGGAAAG	1055
Qy	341	Asp***LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys	360
Db	1056	GACAGACTTGTGGACGAAGTTGACACTCAAAACCTCCTACATTAGAGCAATCGTGAAAG	1115
Qy	361	GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluGluCys	380
Db	1116	GAGACATTCCGCATGCACCCGCCACTCCCAGGGTCAAAGAAAGTCACAGAACAGAGTGT	1175
Qy	381	***IleAsnGly***Val***ProGluGlyAlaLeu*****PheAsnValTrpGlnVal	400
Db	1176	GAGATTAATGGATATGTGATCCCAGAGGGAGCATTGATTCTCTTCAATGTATGGCAAGTA	1235
Qy	401	Gly***Asp***LysTyrTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu	420
Db	1236	GGAAGAGACCCCAAATACTGGGACAGACCATCGGAGTCCGTCCTGAGAGGTTCTAGAG	1295
Qy	421	Thr***AlaGluGlyGluAla***LeuAspLeuArgGly***HisPheGlnLeuLeu	440
Db	1296	ACAGGGGCTGAAGGGGAAGCAGGGCTCTTGATCTTAGGGACAACATTCAACTTCTC	1355
Qy	441	ProPheGlySerGlyArg***MetCysProGlyVal***LeuAlaThrSerGly***Ala	460
Db	1356	CCATTGGGTCTGGGAGGAGAATGTGCCCTGGAGTCATCTGGCTACTCGGGAATGGCA	1415
Qy	461	ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGlyGln	480
Db	1416	ACACTTCTTGCATCTCTTATTCACTGCTTGACTTGCAAGTGCTGGTCCACAAGGACAG	1475
Qy	481	IleLeuLysGly***AspAlaLysValSerMetGluGluArgAlaGlyLeuThrValPro	500
Db	1476	ATATTGAAGGGTGGTGACGCCAAAGTTAGCATGGAAGAGAGGCCGGCTCACTGTTCCA	1535
Qy	501	ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu	520
Db	1536	AGGGCACATAGTCTTGTCTGTGTTCCACTTGCAAGGATGGCGTTGCATCTAAACTCCT	1595
Qy	521	Ser 521	
Db	1596	TCT 1598	

RESULT .10

AF022462

LOCUS AF022462 1824 bp mRNA linear PLN 02-MAR-1999
DEFINITION Glycine max cytochrome P450 monooxygenase CYP93C1p (CYP93C1) mRNA, complete cds.
ACCESSION AF022462
VERSION AF022462.1 GI:2739005
KEYWORDS .
SOURCE Glycine max.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicots; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilioideae; Phaseoleae;
 Glycine.
REFERENCE
AUTHORS 1 (bases 1 to 1824)
 Siminszky, B., Corbin, F.T., Ward, E.R., Fleischmann, T.J. and
 Dewey, R.E.
TITLE Expression of a soybean cytochrome P450 monooxygenase cDNA in yeast
 and tobacco enhances the metabolism of phenylurea herbicides
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (4), 1750-1755 (1999)
MEDLINE 99145622
PUBMED 9990096
REFERENCE 2 (bases 1 to 1824)
AUTHORS Siminszky, B., Dewey, R.E. and Corbin, F.T.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-1997) Crop Science, North Carolina State
 University, Box 7620, Raleigh, NC 27695, USA
FEATURES
 source Location/Qualifiers
 1. .1824
 /organism="Glycine max"
 /db_xref="taxon:3847"
 gene 1. .1824
 /gene="CYP93C1"
 CDS 54. .1619
 /gene="CYP93C1"
 /note="cytochrome P450 monooxygenase"
 /codon_start=1
 /product="CYP93C1p"
 /protein_id="AAB94591.1"
 /db_xref="GI:2739006"
 /translation="MLLELALGLLVLAFLHLRPTPTAKSKALRHLPPSPKPRLPF
 IGHLHLLDKLLHYALIDLSSKKHGPLFSLYFGSMPTVVASTPELFKLFLQTHEATSFN
 TRFQTSAIRRLTYDSSVAMVPFGPYWKFVRKLIMNDLPNATTVNKLRPLRTQQTRKF
 RVMAQGAEAQKPLDLTEELLKWTNSTISMMLGEAAEIRDIAREVLKIFGEYSLTDFI
 WPLKHLVKGVKYEKRIIDILNKFDPVVERVIKKREIVRRRKNGEVVEGEVSGVFLDTL
 LEFAEDETMEIKITKDHIEGLVVDFFSAGTDSTA VATEWALAE LINNPKVLEKAREEV
 YSVVGKDRLVDEVDTQNL PYIRAIKVETFRMHPLPVVKRCTEECEINGYVIPEGAL
 ILFNVWQVGRDPKYWDRPSEFRPERFLETGAEGEAGPLDLRGQHFQLLPFGSGRRMCP
 GVNLATSGMATLLASLIQCFDLQVLGPQGQILKGDDAKVSMEERAGLTVPRAHSLVCV
 PLARIGVASKLLS"
BASE COUNT 497 a 448 c 416 g 463 t

Alignment Scores:

Pred. No.:	1.3e-257	Length:	1824
Score:	2251.00	Matches:	451
Percent Similarity:	86.76%	Conservative:	1
Best Local Similarity:	86.56%	Mismatches:	69
Query Match:	93.95%	Indels:	0
DB:	8	Gaps:	0

US-09-857-581-66 (1-521) x AF022462 (1-1824)

Qy 21 ThrPro***Ala***SerLysAlaLeuArgHisLeuProAsnProProSerPro***Pro 40

Db	114	ACACCCACTGAAAATCAAAGCACTCGCCATCTCCAAACCCACCAAGCCCCAAAGCCT	173
Qy	41	ArgLeuProPheIleGlyHis***HisLeuLeuLysAspLysLeuLeuHisTyrAla***	60
Db	174	CGTCTTCCCTTCATAGGACACCTTCATCTCTAAAAGACAAACTTCTCCACTACGCACTC	233
Qy	61	IleAspLeuSerLysLysHisGlyProLeuPheSer*****PheGlySerMetProThr	80
Db	234	ATCGACCTCTCCAAAAACATGGTCCCTATTCTCTACTTTGGCTCATGCCAAC	293
Qy	81	ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln*****GluAlaThrSer	100
Db	294	GTTGTTGCCCTCACACCAGAATTGTTCAAGCTTCTCCAAACGCACGAGGCAACTTCC	353
Qy	101	Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla	120
Db	354	TTCAACACAAGGTTCCAAACCTCAGCCATAAGACGCCTCACCTATGATAGCTCAGTGGCC	413
Qy	121	*****Pro***GlyProTyrTrp***PheValArgLysLeuIleMetAsnAspLeuLeu	140
Db	414	ATGGTTCCCTTCGGACCTTACTGGAAAGTTCTGAGGAAGCTCATCATGAACGACCTTCCC	473
Qy	141	AsnAlaThrThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgLys***Leu	160
Db	474	AACGCCACCACGTAAACAAGTTGAGGCCTTGAGGACCCAACAGACCCGCAAGTTCCCT	533
Qy	161	Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluGluLeuLeu	180
Db	534	AGGGTTATGGCCAAGGGCGCAGAGGCACAGAACAGCCCCTTGACTTGACCGAGGAGCTCTG	593
Qy	181	LysTrp***AsnSerThr***SerMetMet***LeuGlyGluAlaGluGluIleArgAsp	200
Db	594	AAATGGACCAACAGCACCATCTCCATGATGATGCTGGCGAGGCTGAGGAGATCAGAGAC	653
Qy	201	IleAlaArgGluValLeuLysIle***GlyGluTyrSerLeuThrAspPheIle***Pro	220
Db	654	ATCGCTCGCGAGGTTCTTAAGATCTTGGCGAATACAGCCTCACTGACTCATCTGGCCA	713
Qy	221	LeuLys***LeuLysValGlyLysTyrGluLysArgIleAspAspIleLeuAsnLysPhe	240
Db	714	TTGAAGCATCTCAAGGTTGGAAAGTATGAGAAGAGGATCGACGACATCTGAACAAAGTTC	773
Qy	241	AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn	260
Db	774	GACCCTGTCGTTGAAAGGGTCATCAAGAACGCCGTGAGATCGTGAGGAGAAAGAAC	833
Qy	261	GlyGlu*****GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAla	280
Db	834	GGAGAGGTTGTTGAGGGTGAGGTCAAGGGGGTTTCCTGACACTTGCTGAATTGCT	893
Qy	281	GluAspGluThr***GluIleLysIleThrLys*****IleLysGlyLeuValValAsp	300
Db	894	GAGGATGAGACCATGGAGATCAAATCACCAAGGACCACATCGAGGGTCTGTTGTCGAC	953
Qy	301	***PheSerAlaGly***AspSerThrAla*****ThrGluTrpAlaLeuAlaGluLeu	320
Db	954	TTTTTCTCGGCAGGAACAGACTCCACAGCGGGTGGCAACAGAGTGGGCATTGGCAGAAC	1013

Qy 321 IleAsnAsnPro***ValLeu*****AlaArgGluGlu***TyrSerValValGlyLys 340
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1014 ATCAACAACTCCTAAGGTGTTGGAAAAGGCTCGTGAGGAGGTCTACAGTGTGGAAAG 1073

 Qy 341 Asp***LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1074 GACAGACTTGTGGACGAAGTTGACACTCAAACCTCCTTACATTAGAGCAATCGTGAAG 1133

 Qy 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluGluCys 380
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1134 GAGACATTCCGCATGCACCCGCCACTCCCAGTGGTCAAAAGAAAGTGCACAGAAGAGTGT 1193

 Qy ***IleAsnGly***Val***ProGluGlyAlaLeu*****PheAsnValTrpGlnVal 400
 ||||| ||||| ||||| ||||| ||||| |||||
 Db 1194 GAGATTAATGGATATGTGATCCCAGAGGGAGCATTGATTCTCTCAATGTATGGCAAGTA 1253

 Qy 401 Gly***Asp***LysTyrTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu 420
 ||||| ||||| ||||| ||||| ||||| |||||
 Db 1254 GGAAGAGACCCCCAAATACTGGGACAGACCATCGGAGTTCCGTCTGAGAGGTTCTAGAG 1313

 Qy 421 Thr***AlaGluGlyGluAla*****LeuAspLeuArgGly***HisPheGlnLeuLeu 440
 ||||| ||||| ||||| ||||| ||||| |||||
 Db 1314 ACAGGGGCTGAAGGGGAAGCAGGGCCTCTGATCTTAGGGACAACATTTCAACTTCTC 1373

 Qy 441 ProPheGlySerGlyArg***MetCysProGlyVal***LeuAlaThrSerGly***Ala 460
 ||||| ||||| ||||| ||||| ||||| |||||
 Db 1374 CCATTGGGTCTGGGAGGAGAACATGTGCCCTGGAGTCAATCTGGCTACTCGGGAATGGCA 1433

 Qy 461 ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGlyGln 480
 ||||| ||||| ||||| ||||| ||||| |||||
 Db 1434 ACACTTCTTGCATCTTATTCAAGTGCCTCGACTTGCAAGTGTGGTCCACAAGGACAG 1493

 Qy 481 IleLeuLysGly***AspAlaLysValSerMetGluGluArgAlaGlyLeuThrValPro 500
 ||||| ||||| ||||| ||||| |||||
 Db 1494 ATATTGAAGGGTGGTGACGCCAAAGTTAGCATGGAAGAGAGGCCGGCTCACTGTTCCA 1553

 Qy 501 ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu 520
 ||||| ||||| ||||| ||||| |||||
 Db 1554 AGGGCACATAGTCTTGTCTGTGTTCCACTTGCAAGGATCGCGTTGCATCTAAACTCCTT 1613

 Qy 521 Ser 521
 |||
 Db 1614 TCT 1616

RESULT 11

AF195799

LOCUS AF195799 1824 bp mRNA linear PLN 16-FEB-2000
 DEFINITION Glycine max isoflavone synthase 2 (ifs2) mRNA, complete cds.
 ACCESSION AF195799
 VERSION AF195799.1 GI:6979521
 KEYWORDS
 SOURCE Glycine max.
 ORGANISM Glycine max.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicots; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

REFERENCE 1 (bases 1 to 1824)

AUTHORS Jung, W., Yu, O., Lau, S.M., O'Keefe, D.P., Odell, J., Fader, G. and McGonigle, B.

TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes

JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)

MEDLINE 20124255

PUBMED 10657130

REFERENCE 2 (bases 1 to 1824)

AUTHORS Jung, W., Yu, O., Odell, J., Fader, G. and McGonigle, B.

TITLE Direct Submission

JOURNAL Submitted (18-OCT-1999) Nutrition and Health, DuPont, P.O. Box 80402, Wilmington, DE 19880-0402, USA

FEATURES Location/Qualifiers

source 1. .1824
 /organism="Glycine max"
 /db_xref="taxon:3847"

gene 1. .1824
 /gene="ifs2"

CDS 54. .1619
 /gene="ifs2"
 /codon_start=1
 /product="isoflavone synthase 2"
 /protein_id="AAF34520.1"
 /db_xref="GI:6979522"
 /translation="MLLEALAGLLVLALFLHLRPTPTAKSKALRHLPPSPKPRLPF
 IGHLHLLDKLLHYALIDLSSKKHGPLFSLYFGSMPTVVAStPELFKLFLQTHATSFN
 TRFQTSAIRRLTYDSSVAMVPFGPYWKFRKLIMNDLPNATTVNKLRLRTQQTRKF
 RVMAQGAEAQKPLDLTEELLKWTNSTISMMMLGEAEEIIRDIAREVLIKIFGEYSLTD
 WPLKHLKVGKYEKRIDDILNKFDPVVERVIKKRREIVRRRNGEVVEGEVSGVFLDTL
 LEFAEDETMEIKITKDHIIEGLVVDFFSAGTDSTAVATEWALAEINNPKVLEKAREEV
 YSVVGKDRLVDEVDTQNLPLYIRAIKETFRMHPPPLPVVKRKCTEECEINGYVIPEGAL
 ILFNVWQVGRDPKYWDRPSEFRPERFLETGAEGEAGPLDLRGQHFQLLPFGSGRRMCP
 GVNLATSGMATLLASLIQCFDLQVLGPQQQILKGGDAKVSMEEERAGLTVPRAHSLVCV
 PLARIGVASKLLS"

BASE COUNT 497 a 448 c 416 g 463 t

ORIGIN

Alignment Scores:

Pred. No.:	1.3e-257	Length:	1824
Score:	2251.00	Matches:	451
Percent Similarity:	86.76%	Conservative:	1
Best Local Similarity:	86.56%	Mismatches:	69
Query Match:	93.95%	Indels:	0
DB:	8	Gaps:	0

US-09-857-581-66 (1-521) x AF195799 (1-1824)

Qy	1 MetLeuLeuGluLeuAlaLeuGlyLeu***ValLeuAlaLeuPhe***HisLeuArgPro	20
Db	54 ATGTTGCTTGAACTTGCAC TTGGTTATTGGTTTGCGCTCTGTTCTGCACTTGCCTCCC	113
Qy	21 ThrPro***Ala***SerLysAlaLeuArgHisLeuProAsnProProSerPro***Pro	40
Db	114 ACACCCACTGCAAATCAAAGCACTTCGCCATCTCCCAAACCCACCAAGCCCAAAGCCT	173

Qy 41 ArgLeuProPheIleGlyHis***HisLeuLeuLysAspLysLeuLeuHisTyrAla*** 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 174 CGTCTTCCCTTCATAGGACACCTTCATCTCTAAAGACAAACTCTCCACTACGCACTC 233

 Qy 61 IleAspLeuSerLysLysHisGlyProLeuPheSer*****PheGlySerMetProThr 80
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 234 ATCGACCTCTCCAAAAACATGGTCCCTATTCTCTCTACTTTGGCTCATGCCAACC 293

 Qy 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln*****GluAlaThrSer 100
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 294 GTTGTGCCTCCACACCAGAATTGTTCAAGCTCTCCTCCAAACGCACGAGGCAACTTCC 353

 Qy 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla 120
 ||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 354 TTCAACACAAGGTTCAAACCTCAGCCATAAGACGCCACCTATGATAGCTCAGTGGCC 413

 Qy 121 *****Pro***GlyProTyrTrp***PheValArgLysLeuIleMetAsnAspLeuLeu 140
 ||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 414 ATGGTTCCCTCGGACCTTACTGGAAGTTCGTGAGGAAGCTCATCATGAACGACCTTCCC 473

 Qy 141 AsnAlaThrThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgLys***Leu 160
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 474 AACGCCACCACTGTAAACAAGTTGAGGCCTTGAGGACCCAACAGACCCGCAAGTTCTT 533

 Qy 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluGluLeuLeu 180
 ||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 534 AGGGTTATGCCAACAGCACCCTCCATGATGATGCTCGCGAGGCTGAGGAGATCAGAGAC 593

 Qy 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyGluAlaGluGluIleArgAsp 200
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 594 AAATGGACCAACAGCACCCTCCATGATGATGCTCGCGAGGCTGAGGAGATCAGAGAC 653

 Qy 201 IleAlaArgGluValLeuLysIle***GlyGluTyrSerLeuThrAspPheIle***Pro 220
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 654 ATCGCTCGCGAGGTTCTTAAGATCTTGGCGAATACAGCCTCACTGACTTCATCTGGCCA 713

 Qy 221 LeuLys***LeuLysValGlyLysTyrGluLysArgIleAspAspIleLeuAsnLysPhe 240
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 714 TTGAAGCATCTCAAGTTGGAAAGTATGAGAAGAGGATCGACGACATCTGAACAAGTTC 773

 Qy 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 774 GACCCTGTCGTTGAAAGGGTCATCAAGAACGCGCGTGAGATCGTGAGGAGGAGAAAGAAC 833

 Qy 261 GlyGlu*****GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAla 280
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 834 GGAGAGGTTGTTGAGGGTGAGGTCAAGCAGGACACATCGAGGGTCTGACACTTGCTGAATTGCT 893

 Qy 281 GluAspGluThr***GluIleLysIleThrLys*****IleLysGlyLeuValValAsp 300
 ||||| ||||| ||||| ||||| ||||| ||||| :|||:
 Db 894 GAGGATGAGACCATGGAGATCAAATCACCAAGGACACATCGAGGGTCTGTTGTCGAC 953

 Qy 301 ***PheSerAlaGly***AspSerThrAla*****ThrGluTrpAlaLeuAlaGluLeu 320
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 954 TTTTTCTCGGCAGGAACAGACTCCACAGCGGTGGCAACAGAGTGGCATTGGCAGAACTC 1013

 Qy 321 IleAsnAsnPro***ValLeu*****AlaArgGluGlu***TyrSerValValGlyLys 340

Db	1014	ATCAACAATCCTAAGGTGTTGGAAAAGGCTCGTGAGGAGGTCTACAGTGTGTGGAAAG	1073
Qy	341	Asp***LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys	360
Db	1074	GACAGACTTGTGGACGAAGTTGACACTCAAAACCTCCTTACATTAGAGCAATCGTGAAG	1133
Qy	361	GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluGluCys	380
Db	1134	GAGACATTCCGCATGCACCCGCCACTCCCAGTGGTCAAAAGAAAGTGCACAGAAGAGTGT	1193
Qy	381	***IleAsnGly***Val***ProGluGlyAlaLeu*****PheAsnValTrpGlnVal	400
Db	1194	GAGATTAATGGATATGTGATCCCAGAGGGAGCATTGATTCTCTCAATGTATGGCAAGTA	1253
Qy	401	Gly***Asp***LysTyrTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu	420
Db	1254	GGAAGAGACCCCAAATACTGGGACAGACCATCGGAGTTCCGTCTGAGAGGTTCTAGAG	1313
Qy	421	Thr***AlaGluGlyGluAla*****LeuAspLeuArgGly***HisPheGlnLeuLeu	440
Db	1314	ACAGGGGCTGAAGGGGAAGCAGGGCCTTTGATCTTAGGGAACATTTCAACTTCTC	1373
Qy	441	ProPheGlySerGlyArg***MetCysProGlyVal***LeuAlaThrSerGly***Ala	460
Db	1374	CCATTGGGTCTGGGAGGAGAATGTGCCCTGGAGTCAATCTGGCTACTTCGGGAATGGCA	1433
Qy	461	ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGly	480
Db	1434	ACACTTCTTGCATCTCTTATTCAAGTGCTCGACTTGCAAGTGCTGGTCCACAAGGACAG	1493
Qy	481	IleLeuLysGly***AspAlaLysValSerMetGluGluArgAlaGlyLeuThrValPro	500
Db	1494	ATATTGAAGGGTGGTGACGCCAAAGTTAGCATGGAAGAGAGGCCGGCTCACTGTTCCA	1553
Qy	501	ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu	520
Db	1554	AGGGCACATAGTCTTGTCTGTGTTCCACTTGCAAGGATCGCGTTGCATCTAAACTCCTT	1613
Qy	521	Ser 521	
Db	1614	TCT 1616	

RESULT 12

AF195819

LOCUS AF195819 1902 bp DNA linear PLN 23-MAR-2000
 DEFINITION Glycine max isoflavone synthase 2 (ifs2) gene, complete cds.
 ACCESSION AF195819
 VERSION AF195819.1 GI:7288454
 KEYWORDS .
 SOURCE Glycine max.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicots; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE 1 (bases 1 to 1902)

AUTHORS Jung,W., Yu,O., Lau,S.M., O'Keefe,D.P., Odell,J., Fader,G. and McGonigle,B.
 TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes
 JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)
 MEDLINE 20124255
 PUBMED 10657130
 REFERENCE 2 (bases 1 to 1902)
 AUTHORS Jung,W., Yu,O., Odell,J., Fader,G. and McGonigle,B.
 TITLE Direct Submission
 JOURNAL Submitted (18-OCT-1999) Nutrition and Health, DuPont, PO Box 80402, Wilmington, DE 19880-0402, USA
 FEATURES Location/Qualifiers
 source 1. .1902
 /organism="Glycine max"
 /db_xref="taxon:3847"
 gene 1. .1902
 /gene="ifs2"
 mRNA join(1. .948,1085. .1902)
 /gene="ifs2"
 /product="isoflavone synthase 2"
 CDS join(52. .948,1085. .1753)
 /gene="ifs2"
 /note="cytochrome P450"
 /codon_start=1
 /product="isoflavone synthase 2"
 /protein_id="AAF45143.1"
 /db_xref="GI:7288455"
 /translation="MLLELALGLLVLAFLHRLRPTPTAKSKALRHLNPSPPKPRLPF
 IGHLHLLKDPLLHYALIDLSKKHGPLFSLYFGSMPTVAVSTPELFKLFQLTHEATSFN
 TRFQTSAIRRLTYDSSVAMVPFGPYWKFRKLIMNDLPNATTVNKLRPLRTQQTRKF
 RVMAQGAEAAQKPLDLTEELLKWTNSTISMMMLGEAEEIRDIAREVLIKFGESLTDFI
 WPLKHLVKGYEKIRRDLINKFDPVVERVIKKREIVRRRNGEVVEGEVSGVFLDTL
 LEFAEDETMEIKITKDHIEGLVVDFFSAGTDSTAVATEWALAEINNPKVLEKAREEV-
 YSVVGKDRLVDEVDTQNLPIRAIVKETFRMHPPLPVVKRKCTEECEINGYVPEGAL
 ILFNVWQVGRDPKYWDRPSEFRPERFLETGAEGEAGPLDLRGQHFQLLPGSGRMRCP
 GVNLATSGMATLLASLIQCFDLQVLGPQGQILKGDAKVSMEERAGLTVPRAHSLVCV
 PLARIGVASKLIS"

Alignment Scores:

Pred. No.: 1.06e-251 Length: 1902
Score: 2201.50 Matches: 451
Percent Similarity: 79.72% Conservative: 1
Best Local Similarity: 79.54% Mismatches: 69
Query Match: 91.88% Indels: 46
DB: 8 Gaps: 1

US-09-857-581-66 (1-521) x AF195819 (1-1902)

Qy 21 ThrPro***Ala***SerLysAlaLeuArgHisLeuProAsnProProSerPro***Pro 40
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 112 ACACCCACTGAAAATCAAAGCACTTCGCCATCTCCAAACCCACCAAGCCAAAGCCT 171
 Qy 41 ArgLeuProPheIleGlyHis***HisLeuLeuLysAspLysLeuLeuHisTyrAla*** 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 172 CGTCTTCCCTTCATAGGACACCTTCATCTCTAAAAGACAAACTTCTCCACTACGCAC 231
 Qy 61 IleAspLeuSerLysLysHisGlyProLeuPheSer*****PheGlySerMetProThr 80
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 232 ATCGACCTCTCCAAAAAACATGGTCCCTATTCTCTCTACTTTGGCTCCATGCCAACC 291
 Qy 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln*****GluAlaThrSer 100
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 292 GTTGTGCGCTCCACACCAGAATTGTTCAAGCTCTCCTCCAAACGCACGAGGCAACTTCC 351
 Qy 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 352 TTCAACACAAGGTTCAAACCTCAGCCATAAGACGCCACCTATGATAGCTCAGTGGCC 411
 Qy 121 *****Pro***GlyProTyrTrp***PheValArgLysLeuIleMetAsnAspLeuLeu 140
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 412 ATGGTTCCCTTCGGACCTTACTGGAAGTTCGTGAGGAAGCTCATCATGAACGACCTTCCC 471
 Qy 141 AsnAlaThrThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgLys***Leu 160
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 472 AACGCCACCAACTGTAAACAAGTTGAGGCCTTGAGGACCCAACAGACCCGCAAGTTCC 531
 Qy 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluGluLeuLeu 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 532 AGGGTTATGGCCAAGGCGCAGAGGCACAGAAGCCCCCTGACTGACCGAGGAGCTCTG 591
 Qy 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyGluAlaGluGluIleArgAsp 200
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 592 AAATGGACCAACAGCACCATCTCCATGATGATGCTCGCGAGGCTGAGGAGATCAGAGAC 651
 Qy 201 IleAlaArgGluValLeuLysIle***GlyGluTyrSerLeuThrAspPheIle***Pro 220
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 652 ATCGCTCGCGAGGTTCTTAAGATCTTGGCGAATACAGCCTCACTGACTTCATCTGGCCA 711
 Qy 221 LeuLys***LeuLysValGlyLysTyrGluLysArgIleAspAspIleLeuAsnLysPhe 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 712 TTGAAGCATCTCAAGGTTGGAAAGTATGAGAAGAGGATCGACGACATCTGAACAAGTT 771
 Qy 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 772 GACCCTGTCGTTGAAAGGGTCATCAAGAACGCGCGTGAGGATCGTGGAGGAGAAAGAAC 831
 Qy 261 GlyGlu*****GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAla 280
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 832 GGAGAGGTTGGAGGGTGAGGTCAGCGGGTTTCCTGACACTTGCTGAATTGCGT 891
 Qy 281 GluAspGluThr***GluIleLysIleThrLys*****IleLysGlyLeuValVal--- 299
 ||||| ||||| ||||| ||||| ||||| ||||| : : |||||
 Db 892 GAGGATGAGACCATGGAGATCAAATCACCAAGGACCACATCGAGGGTCTGTTGCGT 951
 Qy 299 ----- 299
 Db 952 AGTTTCCTGCTTCATTGATCGAAATATGCAGTATTTGTTAACAGAGATCGAGAA 1011

Qy 299 -----
 Db 1012 TTGACATTATATTATGTGGTGGCAATTAATTAACGGTACGCATTCTTAATCGATAT 1071
 Qy 300 -----Asp***PheSerAlaGly***AspSerThrAla*****ThrGluTr 315
 ||||| ||||||||| ||||| ||||| ||||| |||||
 Db 1072 TGTGTATGTGCAGGACTTTCTGGCAGGAACAGACTCCACAGCGGTGGCAACAGAGTG 1131
 Qy 315 pAlaLeuAlaGluLeuIleAsnAsnPro***ValLeu*****AlaArgGluGlu***Ty 335
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1132 GGCATTGGCAGAACTCATCAACAATCCTAAGGTGTTGGAAAAGGCTCGTAGGGAGGTCTA 1191
 Qy 335 rSerValValGlyLysAsp***LeuValAspGluValAspThrGlnAsnLeuProTyrIl 355
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1192 CAGTGTGTTGGAAAGGACAGACTTGACGAGTTGACACTCAAAACCTTCCTTACAT 1251
 Qy 355 eArgAlaIleValLysGluThrPheArgMetHisProProLeuProValValLysArgLy 375
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1252 TAGAGCAATCGTGAGGAGACATTCCGCATGCACCCGCCACTCCCAGTGGTCAAAAGAAA 1311
 Qy 375 sCys***GluGluCys***IleAsnGly***Val***ProGluGlyAlaLeu*****Ph 395
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1312 GTGCACAGAACAGACTGTGAGATTATGGATATGTGATCCCAGAGGGAGCATTGATTCTCTT 1371
 Qy 395 eAsnValTrpGlnValGly***Asp***LysTyrTrpAspArgProSerGlu***ArgPr 415
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1372 CAATGTATGGCAAGTAGGAAGAGACCCCCAAACTGGACAGACCATCGGAGTTCCGTCC 1431
 Qy 415 oGluArgPheLeuGluThr***AlaGluGlyGluAla*****LeuAspLeuArgGly** 435
 ||||| ||||| ||||| ||||| ||||| |||||
 Db 1432 TGAGAGGTTCCCTAGAGACAGGGCTGAAGGGGAAGCAGGGCCTTGATCTTAGGGACA 1491
 Qy 435 *HisPheGlnLeuLeuProPheGlySerGlyArg***MetCysProGlyVal***LeuAl 455
 ||||| ||||| ||||| ||||| ||||| |||||
 Db 1492 ACATTTCAACTCTCCCATTGGGTCTGGGAGGAGATGTGCCCTGGAGTCATCTGGC 1551
 Qy 455 aThrSerGly***AlaThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLe 475
 ||||| ||||| ||||| ||||| ||||| |||||
 Db 1552 TACTCGGGAATGGCAACACTTCTGCATCTTATTCAAGTGCTTCGACTTGCAAGTGCT 1611
 Qy 475 uGlyProGlnGlyGlnIleLeuLysGly***AspAlaLysValSerMetGluGluArgAl 495
 ||||| ||||| ||||| ||||| |||||
 Db 1612 GGGTCCACAAGGACAGATATTGAAGGGTGGTGACGCCAAAGTTAGCATGGAAGAGAGAGC 1671
 Qy 495 aGlyLeuThrValProArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyVa 515
 ||||| ||||| ||||| ||||| |||||
 Db 1672 CGGCCTCACTGTTCCAAGGGCACATAGTCTGTGTTCCATTGCAAGGATCGCGT 1731
 Qy 515 lAlaSerLysLeuLeuSer 521
 ||||| |||||
 Db 1732 TGCATCTAAACTCCTTCT 1750

RESULT 13

AF195818

LOCUS AF195818 1800 bp DNA linear PLN 23-MAR-2000
 DEFINITION Glycine max isoflavone synthase 1 (ifs1) gene, partial cds.

ACCESSION AF195818
 VERSION AF195818.1 GI:7288452
 KEYWORDS
 SOURCE Glycine max.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicots; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE 1 (bases 1 to 1800)
 AUTHORS Jung,W., Yu,O., Lau,S.M., O'Keefe,D.P., Odell,J., Fader,G. and
 McGonigle,B.
 TITLE Identification and expression of isoflavone synthase, the key
 enzyme for biosynthesis of isoflavones in legumes
 JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)
 MEDLINE 20124255
 PUBMED 10657130
 REFERENCE 2 (bases 1 to 1800)
 AUTHORS Jung,W., Yu,O., Odell,J., Fader,G. and McGonigle,B.
 TITLE Direct Submission
 JOURNAL Submitted (18-OCT-1999) Nutrition and Health, DuPont, PO Box 80402,
 Wilmington, DE 19880-0402, USA
 FEATURES Location/Qualifiers
 source 1. .1800
 /organism="Glycine max"
 /db_xref="taxon:3847"
 gene <1. .>1800
 /gene="ifs1"
 mRNA join(<1. .893,1112. .>1800)
 /gene="ifs1"
 /product="isoflavone synthase 1"
 CDS join(<1. .893,1112. .1780)
 /gene="ifs1"
 /note="cytochrome P450"
 /codon_start=3
 /product="isoflavone synthase 1"
 /protein_id="AAF45142.1"
 /db_xref="GI:7288453"
 /translation="LELALGLFVLALFLHLRPTPSAKSKALRHLNPSPPKPRLPFIG
 HLHLLKDKLHYALIDLSSKKHGPLFSLSGSMPTVVASTPELFKLFLQTHEATSFNTR
 FQTSAIRRRTYDNSVAMVPFGPYWKFVRKLIMNDLLNATTVNKLRPLRTQQIRKFLRV
 MAQSAEAQKPLDVTEELLKWTNSTISMMMLGEAEEIRDIAREVLKIFGEYSLTDFIWP
 LKYLKGVKYEKRIDDILNKFDPPVERVIKKRREIVRRRNQEVVEGEASGVFLDTLLE
 FAEDETMEIKITKEQIKGLVVDFFSAGTDSTAVATEWALAEINNPRVLQKAREEVYS
 VVGKDRLVDEVDTQNLPYIRAIKVETFRMHPPPLPVVKRKCTEECEINGYVPEGALVL
 FNVWQVGRDPKYWDRPSEFRPERFLETGAEGEAGPLDLRGQHFQLLPFGSGRRMCPGV
 NLATSGMATLLASLIQCFDLQVLGPQGQILKGDDAKVSMEERAGLTVPRAHSLVCVPL
 ARIGVASKLLS".
 BASE COUNT 476 a 441 c 413 g 470 t
 ORIGIN

Alignment Scores:

Pred. No.:	3.04e-250	Length:	1800
Score:	2189.00	Matches:	451
Percent Similarity:	76.18%	Conservative:	0
Best Local Similarity:	76.18%	Mismatches:	68
Query Match:	91.36%	Indels:	74

DB:

8

Gaps:

1

US-09-857-581-66 (1-521) x AF195818 (1-1800)

Qy	3	LeuGluLeuAlaLeuGlyLeu***ValLeuAlaLeuPhe***HisLeuArgProThrPro	22
Db	3	CTGGAACTTGCACTTGGTTGTTGTAGCTTCTGCACTTGCCTCGTCCCACACCA	62
Qy	23	***Ala***SerLysAlaLeuArgHisLeuProAsnProProSerPro***ProArgLeu	42
Db	63	AGTGCAAAATCAAAGCACTTCGCCACCTCCAAACCCCTCCAAGGCCAAAGCCTCGTCTT	122
Qy	43	ProPheIleGlyHis***HisLeuLeuLysAspLysLeuLeuHisTyrAla***IleAsp	62
Db	123	CCCTTCATTGGCCACCTCACCTCTAAAAGATAAACTCTCCACTATGCACTCATCGAT	182
Qy	63	LeuSerLysLysHisGlyProLeuPheSer*****PheGlySerMetProThrValVal	82
Db	183	CTCTCCAAAAGCATGGCCCTTATTCTCTCTCCCTCGGCTCCATGCCAACCGTCGTT	242
Qy	83	AlaSerThrProGluLeuPheLysLeuPheLeuGln*****GluAlaThrSerPhe***	102
Db	243	GCCTCCACCCCTGAGTTGTTCAAGCTCTCCCTCAAACCCACGAGGCAACTCCTTCAAC	302
Qy	103	ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla*****	122
Db	303	ACAAGGTTCCAAACCTCTGCCATAAGACGCCACTTACGACAACACTGTGGCCATGGTT	362
Qy	123	Pro***GlyProTyrTrp***PheValArgLysLeuIleMetAsnAspLeuLeuAsnAla	142
Db	363	CCATTGGACCTTACTGGAAGTTCTCGTGAGGAAGCTCATCATGAACGACCTCTAACGCC	422
Qy	143	ThrThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgLys***LeuArg***	162
Db	423	ACCACCGTCAACAAGCTCAGGCCCTTGAGGACCCACAGATCCGAAGTTCCTTAGGGTT	482
Qy	163	MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluGluLeuLeuLysTrp	182
Db	483	ATGGCCCAAAGCGCAGAGGCCAGAAGCCCTTGACGTCACCGAGGAGCTCTCAAATGG	542
Qy	183	***AsnSerThr***SerMetMet***LeuGlyGluAlaGluGluIleArgAspIleAla	202
Db	543	ACCAACAGCACCATCTCATGATGATGCTCGCGAGGCTGAGGAGATCAGAGACATCGCT	602
Qy	203	ArgGluValLeuLysIle***GlyGluTyrSerLeuThrAspPheIle***ProLeuLys	222
Db	603	CGCGAGGTTCTTAAGATCTCGGCGAATACAGCCTCACTGACTCATCTGGCCTTGAAAG	662
Qy	223	***LeuLysValGlyLysTyrGluLysArgIleAspAspIleLeuAsnLysPheAspPro	242
Db	663	TATCTCAAGGTTGGAAAGTATGAGAAGAGGATTGATGACATCTGAACAAGTTCGACCC	722
Qy	243	ValValGluArgValIleLysLysArgArg***IleValArgArgArg***AsnGlyGlu	262
Db	723	GTCTGTTGAAAGGGTCATCAAGAACGCCGTGAGATCGTCAGAAGGAGAAAGAACGGAGAA	782
Qy	263	*****GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAlaGluAsp	282

Db	783	GTTGTTGAGGGCGAGGCCAGCGCGTCTTCCTCGACACTTGCTTGAATTGCTGAGGAC	842
Qy	283	GluThr***GluIleLysIleThrLys*****IleLysGlyLeuValVal-----	299
Db	843	GAGACCATGGAGATCAAAATTACCAAGGAGCAAATCAAGGCCTGTTGT-CGTAAGTTT	901
Qy	299	-----	299
Db	902	CCTTCTTCTCCTACTTATTACTTCTTCATTCATCATATGTATTGGCATTAAATAG	961
Qy	299	-----	299
Db	962	TATACTATATGAGAAAATATGTTACGCACTCACGGTGTAAAGATATGTGGTGTAAAAA	1021
Qy	299	-----	299
Db	1022	AAAAGAGATACAGAACAGTTGCTTTATGCATGTATGTTAACGTATTTACTCAAGTGGAA	1081
Qy	300	-----Asp***PheSerAlaGly***AspSerThr	309
Db	1082	ACTAATTAATTCTCAATTGGGTATGTAGGACTTTCTCTGCAGGGACAGATTCCACA	1141
Qy	310	Ala*****ThrGluTrpAlaLeuAlaGluLeuIleAsnAsnPro***ValLeu*****	329
Db	1142	GCGGTGGCAACAGAGTGGCATTGGCAGAGCTCATCAACAATCCCAGGGTGTGAAAAG	1201
Qy	330	AlaArgGluGlu***TyrSerValValGlyLysAsp***LeuValAspGluValAspThr	349
Db	1202	GCTCGTGAGGAGGTCTACAGTGTGTTGGCAAAGATAGACTCGTTGACGAAGTTGACACT	1261
Qy	350	GlnAsnLeuProTyrIleArgAlaIleValLysGluThrPheArgMetHisProProLeu	369
Db	1262	CAAAACCTTCCTTACATTAGGCCATTGTGAAGGAGACATTCCGAATGCACCCACCACTC	1321
Qy	370	ProValValLysArgLysCys***GluGluCys***IleAsnGly***Val***ProGlu	389
Db	1322	CCAGTGGTCAAAAGAAAGTGCACAGAAGAGTGTGAGATTATGGGTATGTGATCCCAGAG	1381
Qy	390	GlyAlaLeu*****PheAsnValTrpGlnValGly***Asp***LysTyrTrpAspArg	409
Db	1382	GGAGCATTGGTTCTTCATGTTGGCAAGTAGGAAGGGACCCCAAATACTGGGACAGA	1441
Qy	410	ProSerGlu***ArgProGluArgPheLeuGluThr***AlaGluGlyGluAla*****	429
Db	1442	CCATCAGAATTCCGCCCCGAGAGGTTCTTAGAAACTGGTGTGAAGGGAAAGCAGGGCCT	1501
Qy	430	LeuAspLeuArgGly***HisPheGlnLeuLeuProPheGlySerGlyArg***MetCys	449
Db	1502	CTTGATCTTAGGGGCCAGCATTCCAACCTCTCCCATTGGGTCTGGGAGGAGAATGTGC	1561
Qy	450	ProGlyVal***LeuAlaThrSerGly***AlaThrLeuLeuAlaSerLeuIleGlnCys	469
Db	1562	CCTGGTGTCAATTGGCTACTTCAGGAATGGCACACTTCTTGATCTCTATCCAATGC	1621
Qy	470	PheAspLeuGlnValLeuGlyProGlnGlyGlnIleLeuLysGly***AspAlaLysVal	489

Db 1622 TTTGACCTGCAAGTGCTGGCCCTAAGGACAAATATTGAAAGGTGATGCCAAAGTT 1681
 Qy 490 SerMetGluGluArgAlaGlyLeuThrValProArgAlaHisSerLeuValCysValPro 509
 |||||||
 Db 1682 AGCATGGAAAGAGAGAGCTGGCCTCACAGTTCCAAGGGCACATAGTCTCGTTGTGTTCCA 1741
 Qy 510 LeuAlaArgIleGlyValAlaSerLysLeuLeuSer 521
 |||||||
 Db 1742 CTTGCAAGGGATCGCGTTGCATCTAAACTCCTTTCT 1777

RESULT 14
AF195801
LOCUS AF195801 1501 bp mRNA linear PLN 16-FEB-2000
DEFINITION Medicago sativa isoflavone synthase 2 (ifs2) mRNA, partial cds.
ACCESSION AF195801
VERSION AF195801.1 GI:6979525
KEYWORDS
SOURCE Medicago sativa.
ORGANISM Medicago sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 1501)
AUTHORS Jung,W., Yu,O., Lau,S.M., O'Keefe,D.P., Odell,J., Fader,G. and McGonigle,B.
TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes
JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)
MEDLINE 20124255
PUBMED 10657130
REFERENCE 2 (bases 1 to 1501)
AUTHORS Jung,W., Yu,O., Odell,J., Fader,G. and McGonigle,B.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-1999) Nutrition and Health, DuPont, P.O. Box 80402, Wilmington, DE 19880-0402, USA
FEATURES Location/Qualifiers
source 1. .1501
/organism="Medicago sativa"
/db_xref="taxon:3879"
gene <1. .>1501
/gene="ifs2"
CDS <1. .>1501
/gene="ifs2"
/codon_start=3
/product="isoflavone synthase 2"
/protein_id="AAF34522.1"
/db_xref="GI:6979526"
/translation="FLHLRPTPTAKSKALRHLNPSPKPRLPFIGHLHLLDKDLLHY
ALIDLSSKKHGPLFSLYFGSMPTVVASTPELFKLFLQTHEATSFNTRFQTSAIRRLTYD
SSVAMAPFGPYWKFRKLIMNDLLNATTVNKLRPLRTQQIRKFRLRVMAQGAEAQKPLD
LTEELLKWTNSTSMMMLGEAEEIRDIAREVLKIFGEYSLTDFIRPLKHLKVGYEKR
IDDILNKFDPVVERVIKKREIVRRKNGEVVEGEVSGVFLDTLLEFAEDETTEIKIT
KDHIGLVLVDFFSAGTDSTAVATEWALAEINNPKVLEKAREEVYSSVGKDRLVDEVD
TQNL PYIRAIVKETFRMHPPLPVVKRKCTEECEINGYVIPEGALILFNVWQVGRDSKY
WDRPSEFRPERFLETGAEGEARPLDLRGQHFQLLPFGSGRRMCPGVNLATSGMATLLA
SLIQCFDLQVLGPQGQILKGDAKVSMEERAGLTVPRAHSLVCVPLARIG"